XII, No. 3
1961

CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH

DESK

IN THIS ISSUE:

- Classroom Studies Using Teaching Machines
- Students' Values: Emergent or Traditional
- The Diagnostic Use of WISC Subtest Patterns
- Cooperative Study of Educational Problems
- Study of Children's Play in America and Japan

published by
CALIFORNIA TEACHERS ASSOCIATION
1705 Murchison Drive Burlingame California

CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH

Published by the CALIFORNIA TEACHERS ASSOCIATION for the CALIFORNIA ADVISORY COUNCIL ON EDUCATIONAL RESEARCH

Editor

GARFORD G. GORDON
Research Executive, C.T.A.
and Chairman of the Council

Associate Editor

JOHN H. BRIGHT
Assistant Research Executive, C.T.A.
and Secretary of the Council

Editorial Assistant

NANCY J. KOEHLER

●

EDITORIAL BOARD: MEMBERS OF THE COUNCIL

HUGH M. BELL
Professor of Psychology
Chico State College

MERLE H. ELLIOTT
Director of Research
Oakland City Schools

HOWARD A. BOWMAN
Director, Evaluation and Research
Los Angeles City Schools

D. WELTY LEFEVER
Professor of Education
University of Southern California

WILLIAM S. BRISCOE
Professor of Education
University of California, Los Angeles

HENRY W. MAGNUSON
Chief, Bureau of Education Research
State Department of Education

HAROLD D. CARTER
Professor of Education
University of California, Berkeley

HARRY W. SMALLENBURG
Director, Division of Research and Guidance
Los Angeles County Schools

ARTHUR P. COLADARCI
Professor of Education and Psychology
Stanford University

WILLIAM H. STEGEMAN
Director of Research
San Diego City Schools

JAMES CORSON
Executive Secretary
California Association of School Administrators
Burlingame, California

HAROLD WEEKS
Director, Bureau of Research
San Francisco Unified School District

HENRY I. WEITZEL
Research Director
Pasadena City Schools

●

Published at Burlingame five times a year: January, March, May, September, and November. Subscription price: $6.00 per year; single copies, $1.50. All correspondence should be directed to: CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH, 1703 Murchison Drive, Burlingame, California. Checks should be made payable to the CALIFORNIA TEACHERS ASSOCIATION.

Certain articles in this *Journal* are listed in Psychological Abstracts.

Member of the Educational Press Association of America

Entered as second class matter at the Post Office at Burlingame, California

Copyright 1961 by the California Teachers Association

CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH

Volume XII MAY 1961 Number 3

CONTENTS

Planning for Tomorrow, *An Editorial* 98

Teaching Machines: What Have Studies in the Classroom Shown? . . 99
Stanley S. Blank

The Diagnostic Use of WISC Subtest Patterns 116
Kenneth D. Hopkins and William B. Michael

Some Participant Reactions to a Cooperative Approach to the Study of Educational Problems 118
Caseel D. Burke

A Comparative Study of Children's Play in America and Japan . . . 124
May V. Seagoe and Ken Murakami

Meet Your CACER 131

High School Students' Values: Emergent Or Traditional 132
O. E. Thompson

THE EDITORS SAY:

Planning For Tomorrow

The Governor's Committee on Education for the Commonwealth of Pennsylvania has just released its final report. The 139-page document is a provocative statement which discusses the educational problems which a state faces. It is interesting for Californians to note that the critical issues which confront us also are being considered elsewhere in the United States. Such items as guidance, testing, emotionally disturbed children, the gifted child, higher education, scholarships, teacher training, certification, teacher salaries, teaching conditions, organization of school districts, school finance, and many others, are equally as difficult to solve in Pennsylvania as they are in the Golden State.

There is a section, however, on educational research and development which warrants our attention. The report points out that the state educational program is presently in a deficit position, and ways and means are proposed to make up this deficit. But what of 1965, 1970, or 1980? Alert industrial firms do not ask, what of 1965? They plan for tomorrow with today's resources, through research and development programs. No industry worth its salt is without such programs, and they often account for 4 or 5 per cent of current expenditures. In addition, they receive the attention of top-line executives.

The report urges the creation of a special research and development fund to finance state-guided but locally conducted experimental programs on such things as team teaching, track systems, use of educational television and "teaching machines." In other words, the proposal is for research into basic techniques of learning and teaching with a better educational program being the eventual result. The research program should be a continuous, relentless, probing process.

To initiate this research and development program, the Pennsylvania Legislature is being asked to appropriate one-half million dollars. A state governing board would then be empowered to review, approve, and finance locally proposed research, and, of course, initiate its own projects. All results would be published and disseminated among all school districts.

Admittedly, $500,000 is a *David* when placed beside the *Goliath* of total educational expenditures. But, it is an excellent start and one which should be considered by every state—including California.—JHB

Teaching Machines: What Have Studies In The Classroom Shown?

STANLEY S. BLANK

This paper reviews all the studies in published literature to date that are concerned with the use of mechanical devices for individual self-instruction in the classroom.

Despite great variation in the complexity of these devices, they all represent some form of what can be called the tutorial or Socratic method of teaching, in that they present the student with programs of questions and answers, problems to be solved, or exercises to be performed. In addition, they also provide some type of automatic feedback or correction to the student so that he is immediately aware of his progress at each step and is given a basis for correcting his errors. They differ from most other audio-visual media as ordinarily utilized because of several important properties. As defined by Stolurow (22), these properties are:

1. A teaching machine presents information either as statements accompanied by questions, or as paired sets of abstract symbols to be associated together with incomplete associations for testing purposes.

2. A teaching machine requires a response that is overt and measurable.

3. A teaching machine compares, or permits the learner to compare, his response to a predetermined correct response.

4. A teaching machine provides knowledge of results and has the potential for doing this after every response.

5. A teaching machine programs information according to a set of rules that are capable of being made explicit.

6. A teaching machine allows for variability in the presentation rate of information.

7. A teaching machine combines learning and its measurement into a single set of coordinated operations.

8. A teaching machine automatically rejects the correctly-responded-to

---

Stanley S. Blank is a graduate student at the University of California, Berkeley. He obtained his B.A. degree and his teacher training at the University of British Columbia, as well as his M.A. degree in 1960. Mr. Blank taught mathematics for two years in a junior-senior high school in Vancouver, British Columbia.

items so that the learner sees a smaller number of items from trial to trial.

The devices thus represent a way of providing a pre-programmed study-practice combination which simulates the functions of a private tutor in recitation and practice, with immediate correction of errors and feedback to the student.

Early Investigations

The first "teaching machine" appeared in 1926 when Sidney L. Pressey described a "simple device which gives tests and scores and teaches." In this paper Pressey (17) described a machine which presented questions to students who indicated their answers by pushing one of four buttons on the machine. A device on the machine automatically counted the number of correct responses.

The device could also be used for teaching, since it could be mechanically arranged so that the following question would not appear until the correct response was given to the present question.

As Pressey pointed out, a major benefit of the device was that it informed the student immediately if he was right or wrong.

> *"The somewhat astounding way in which the functioning of the apparatus seems to fit in with the so-called 'laws of learning' deserves mention in this connection. The 'law of recency' operates to establish the correct answer in the mind of the subject, since it is always the* LAST *answer which is the right one. The 'law of frequency' also cooperates; by chance, the right response tends to be made most often, since it is the only response by which the subject can go on to the next question."* (17)

The following year Pressey (18) described a machine *"a bit more modern than the previous model."* This machine, similar in appearance to the previous model, was so devised that it would not proceed until the correct response to the presented question was given. Also, the number of times the subject went through the total series of statements and questions making correct responses could be controlled. Thus, after a subject had answered a question correctly the stipulated number of times, the question was automatically passed over in subsequent presentations of the series. In this way the learner's progress was made evident for him by progressive elimination of items.

No experimental studies using these devices were reported, though in a third paper Pressey (19) evaluated what he considered the future possibilities of work along these lines. In his conclusion to this article, Pressey stated: *"Education is a large scale industry, it should use quantity production methods . . ."* and *"The advance of a science is closely dependent upon the development of instruments specifically for the very extensive yet analytical*

research typical both of modern educational investigation and also more generally of the social sciences. New instruments and materials greatly facilitating research may soon appear. There may then be sweeping research advances in this field."

Though Pressey dropped further work on these problems, one of his students, J. K. Little (12) reported a study wherein he attempted to teach a course in educational psychology by machine.

In this experiment, Little used the machine described earlier by Pressey (18). By means of this "drill machine" the student was presented a series of questions (multiple choice or true and false). If the answer was correct, the machine turned up the next question; if incorrect, the same question remained before the student and he was required to find the correct answer before proceeding. A counter dial counted the total number of trials made in completing the test. The students were expected to go through each test in this manner until they did so with no errors.

The investigation was carried on during the winter and spring quarters and the material dealt with covered the regular work in an educational psychology course. Each subject in the experimental group was matched up with one in the control group, which covered the course in the usual manner. The experimental subject and his control were matched on the basis of scores on a pre-test and on the results of an intelligence test.

The procedure followed was for each group to cover a similar amount of work within a given period of time. Both groups were then tested with the same examinations.

As crude as the methods used in this early experiment were, and even though Little lacked today's knowledge of programming, the study still showed surprising results. It was found that approximately twice as great a percentage of the students in the control group received D's and E's as in the experimental group, and fewer made A's and B's.

This experiment is now, of course, mainly of historical interest since recent investigations have provided better methods of control, reinforcement, and programming.

Skinner's Revival of Teaching Machines

Following this study, little more appeared regarding teaching machines until B. F. Skinner's *The Science of Learning and the Art of Teaching* in 1954, wherein he advocated the possibility of more direct educational application of concepts developed in the experimental laboratory. He proposed that his extensive work on the experimental analysis of behavior had direct implications for the teaching process and that its application could be effectively implemented by appropriate instrumentation.

Skinner reviewed the situation as he felt it existed in our schools, pointing out that the child is behaving primarily to escape from the threat of

a series of minor aversive events. In this situation, getting the right answer is in itself an insignificant event, any effect of which is lost amid the anxieties, boredom and aggression which he claimed are the inevitable by-products of aversive control; and that whatever reinforcements are forthcoming in the present-day classrooms are usually accorded by the teacher under conditions far from optimal. Further, he felt, there is a relative infrequency of reinforcement of the child as compared to what empirical evidence showed is required in complex learning situations.

Skinner stated that the advances made in our control of the learning process would suggest a thorough revision of classroom practices utilizing the knowledge gained from these advances which have shown us that the contingencies of reinforcement which would be most efficient in controlling the student cannot be arranged through the personal mediation of the teacher. Therefore, mechanical help is needed. *"The simple fact is that, as a mere reinforcing mechanism, the teacher is out of date. . . . If the teacher is to take advantage of recent advances in the study of learning she must have the help of mechanical devices."* (20) The advantages put forth for such devices were that:

a. Reinforcement for the right answer is immediate;

b. Each child may progress at his own rate;

c. The device makes possible a step-wise presentation through which the most efficient progress to an eventually complex repertoire can be made.

Following Skinner's lead, a number of studies dealing with mechanical teaching devices appeared. Many of these studies were concerned with training problems in the military service and industry. Comparatively few studies dealt with the effectiveness of such devices in the classroom. The latter are reviewed in this paper.

Effectiveness in Teaching Subject Matter

GERMAN LANGUAGE

One of the first attempts to utilize Skinner's principles of programmed learning that appears in the literature was carried out by Ferster and Sapon (9). This was an attempt to teach a first-semester college course in German via teaching machines.

The instructional material used was composed of sheets of paper upon which were written, one beneath the other, pairs of sentences equivalent in German and English. Covering each sheet was a mask which permitted exposure of one line of material at a time, as well as the scoring column which was always visible. The students were required to expose an English

sentence and to write its equivalent in German. The mask was then moved up, exposing the correct translation. If what the student had written and the German conformed exactly, the student took credit for a correct response and passed on to the next item. If an error was made, the student re-covered the German and repeated the process. At the completion of a unit of work, the student was instructed to repeat the unit. This was done until every English sentence had been correctly translated into German in two consecutive attempts. By scanning the scoring columns, the student determined which items could be passed over as they reached the required criterion.

According to the authors, *"the conformance of the student's written German with the materials he uncovers on the second line is the reinforcement that maintains the study behavior."*

The material was administered to 28 adult subjects, none of whom had had any prior German training. The subjects were given three aptitude tests for foreign language learning.

Of the 28 subjects, six finished the lessons and were tested in terms of achievement on content tests. The authors indicated that with a mean time of 47.5 hours, the six subjects learned an amount of German comparable to that presented in a first-semester course. They also concluded that the vocabulary and syntax learned were active in contrast to the largely "passive" vocabulary acquired in a conventional course. It should be noted, however, that this experimental course deleted supplementary reading materials and hence the net recognition vocabulary was smaller.

While Ferster and Sapon must be commended in that they attempted to apply psychological theory to a concrete teaching problem, there appear to be certain methodological inadequacies in the study. The authors, though providing much detail as to means and ranges, provide no statistical interpretation of the control groups. The authors are not able to provide such comparison since the findings of the study are not based on standardized tests, but rather on tests specifically designed to measure the content of the course. These, they felt, would have greater specific validity than currently used standardized tests since these *"attempt to match current textbook and classroom materials and procedures."* However, as Brown and Hodgkinson (3) note, *"specific validity is of little value as opposed to that obtained from standardized tests because the study is an attempt to compare the new procedure with the more conventional procedure. In order to compare the two methods, the evaluative instruments must have general validity for the measurement of achievement under both methods."*

Since no controls were used, the authors' conclusion that *"the vocabulary and syntax learned were more active than that learned in conventional courses"* is questionable.

Finally, the conclusion of the authors that *"the results show the possibility of teaching basic foreign language skills with high efficiency and economy of time for both student and teacher"* seems unwarranted in light

of the lack of comparison with groups taught by conventional methods.

ARITHMETIC

The next reported study that utilizes teaching devices in a class setting was an attempt by Keislar (11) to develop understanding in arithmetic via teaching machines.

The problem in this study was to investigate the possibility of using a multiple-choice method for the automated teaching of "understanding" of areas of triangles. By "understanding" Keislar meant the ability to answer a variety of questions different from those encountered during training, but belonging to the same general class. The broader this class is, the greater is the understanding.

The teaching machine used in this study was an extensive adaption of the Film Rater used in the Navy. Multiple-choice items in a kodachrome film-strip were projected in sequence upon a viewing plate. The learner responded to each item by pressing one of five buttons. If the answer was correct, a green light was turned on and the next item could then be brought into view by pressing a special button. If the answer was wrong, a red light came on; only after turning off this red light could the learner try again. To proceed to the next item, the learner had to answer correctly the item upon which he was working. A special device recorded a graph of all right and wrong answers for each item.

The total program consisted of 120 items, ten of which instructed the learner how to operate the machine and informed him of the goal to be obtained. The balance of the items provided a sequence beginning with concepts of squares, rectangles, their length, width and area, and provided practical problems of adding and subtracting areas and finding lengths and widths of rectangles. The items were so arranged as to provide a "minimum difficulty level" from item to item for each step.

Twenty-eight subjects were used in the study; 14 experimental subjects and their 14 controls. These were matched on the basis of intelligence, sex, reading ability, and pre-test scores. All subjects were selected from the fifth and low-sixth grades.

Both the pre-test and the post-test were of the free-answer type. The questions on the post-test consisted of the same problems as the pre-test, plus another equal number, most of which were more difficult.

The procedure followed consisted of the experimental subjects operating the machine for two to three periods on successive days. The total time spent with the machine ranged from one and a half to two hours. No special instruction was given the control subjects who were used to control for the effects of incidental learning. The post-test was given the day following the end of the machine instruction to each experimental subject and his control.

The result of the post-test was a mean score for the experimental group

of 12.4 with a standard deviation of 5.6. The corresponding control group mean was 5.4 with a standard deviation of 3.7. The results of a sign test showed that all except one of the experimental subjects achieved a higher post-test score than did their matched controls. This difference was significant at the .01 level. An analysis of the post-test answers revealed that the experimental subjects learned less than the author expected they would. On the basis of his analysis of scores and answers, the author concluded that the program was more appropriate for the brighter children. He further concluded that the use of multiple-choice items in automated teaching appears to have some effectiveness under the conditions used in this study.

Several points would lead one to question the conclusions Keislar draws from this study.

Since the control group received no instruction of any kind in the area concerned, one wonders what Keislar means by "some effectiveness." Since no comparisons are possible, it is impossible to judge how effective his instructional method actually was. In this regard, it would have been most interesting had Keislar provided a third group of subjects taught the same unit by conventional methods. In this way he might perhaps have been able to draw some conclusions as to the effectiveness of the machine method as compared to these more conventional methods.

In his discussion of the principles upon which his program is based, the author stated that the student was informed only if he was right or wrong, conceded that the absence of an explanation was likely to heighten the aversive consequences of failure, and hence concluded that it would have been best to adopt a minimum difficulty level for each step so as to avoid, where possible, the students' making errors. However, he neglects to define what he means by "minimum difficulty level," nor does he show how he arrived at this minimum difficulty level. Further, several studies, Coulson (4), Amsel (1), Melaragno (14), would suggest that some negative reinforcement appearing within a sequence of instructional items, all of which are positively reinforcing, will not depress a student's learning of the subject matter. Thus one wonders if the author has demonstrated that this method of instruction has value for teaching any but what he terms "bright" children.

Also, it would seem that the small sample size (14 subjects) would render the generalizing of results to a larger group somewhat questionable.

SPELLING

Another study reporting the effects of teaching-machine instruction in the classroom was carried out by Porter (16). This investigation was carried out during 22 of the normal 32 weeks of spelling instruction given to groups of second-grade and sixth-grade level students in an elementary school. The author described the teaching machine as being of the maximum simplicity.

The teaching materials were fed past two windows in the machine by means of a uni-directional friction-drive roller. The students' responses were written directly on the teaching materials through the lower window. Operation of a lever moved the written response to the upper window, under glass, and exposed the correct response in the lower window. The student was thus able to compare his answer and score himself right or wrong. The operating lever could then be moved again, bringing up the next item. As the author stated, *"the sole mechanical functions of the machine are to provide immediate response confirmation, preventing cheating and keeping subjects from having access to any but the current teaching item."*

An interesting aspect of the entire program was that no spoken instruction was given to students; all necessary instructions were included in the program. The material was programmed to require identification of the spelling words in a meaningful sentence context, matching the correct word to a given definition, matching several words on the basis of letter structure, and writing missing letters in words presented in sentence context.

The subjects in the experimental groups were matched with control subjects on the basis of achievement in other areas and intelligence scores. Both the experimental and control groups were taught exactly the same words.

In his assessment of the experimental findings, Porter stated that on both the second- and sixth-grade levels, spelling achievement, as measured by standardized achievement tests, was significantly superior for the experimental groups.

An interesting outcome, which provides possible implications for the use of teaching machines in teaching atypical students, was that there was no significant relationship between intelligence scores and achievement in the experimental groups, but there was a significant positive relationship in the control group. The scores of weekly tests for the first and second half of the machine lessons were compared and revealed no apparent differences. This finding would seem to support Skinner's contention that the machines themselves would provide some of the reinforcement necessary to keep the level of work at least steady. Another interesting finding of this study was that in the sixth-grade experimental group there was a positive relationship between the number of responses per lesson required of subjects and the consequent achievement, but no relationship between the number of errors per lesson and achievement. This would seem to again raise the question as to whether the steps in programming must be made small as possible so as to avoid the occurrence of errors. Finally, Porter found that 20 per cent of the second graders and 10 per cent of the sixth graders could not match their answers with the answers given in the machine. If this finding is borne out by future research, it may well raise the question as to whether teaching machines should be used with all students.

Effects of Independent Variables

While the prime purpose of this paper is to review those studies that have investigated the over-all effectiveness of teaching-machine instruction in the classroom, it is felt that studies that have sought to find the most effective type of program should also be included.

RESPONSE MODES

The first investigation of this type to appear in the literature is that carried out by Coulson and Silberman (5). The purpose of this study was to investigate three variables:

a. To compare, for teaching effectiveness, the multiple-choice and constructed-response modes.

b. To evaluate the extent of the purported superiority of the small-step item construction.

c. To compare a pre-determined item sequence with a more adaptive procedure (branching), in which a student who had already learned a concept was allowed to skip certain other items covering the same concept. The authors suggest that merely being shown the correct answer to an item after several tries does not insure that the student has actually acquired the necessary knowledge; that if the student does not understand the concept, he should not be presented with more complex questions based upon the original concept.

The experiment consisted of three phases:

a. A training session for each of 80 experimental subjects with a teaching machine operating under one of the eight different teaching procedures being compared.

b. A criterion test given to each subject immediately following the training session to determine how much the subject learned in the training session.

c. The same criterion test given approximately three weeks after the training session to the 80 subjects, and also to a control group of 104 students who were in the same classes with the subjects, but who did not receive teaching machine training.

This last phase was intended to provide a measure of retention in the subjects as well as a comparison of subjects with other students who received no training on the concepts covered in the experiment. From such a comparison the authors thought it possible to determine whether significant learning had occurred in the experimental group as a whole.

The experimental subjects were given three weeks' training with each subject being run under one of the following conditions:

1. Multiple choice, small steps, no branching.
2. Constructed response, small steps, no branching.
3. Multiple choice, large steps, no branching.
4. Constructed response, large steps, no branching.
5. Multiple choice, small steps, branching.
6. Constructed response, small steps, branching.
7. Multiple choice, large steps, branching.
8. Constructed response, large steps, branching.

Ten subjects were randomly assigned to each of the eight experimental conditions. Performance measures of: (a) time required to finish training; (b) score on the multiple choice portion of the criterion test; and (c) score on the constructed response portion of the criterion test were recorded for each subject regardless of mode of instruction. All subjects, experimental and control, were taken from beginning psychology classes in Santa Monica City College. The material covered in the experiment was based on an *elementary psychology course* given at Harvard University.

In their analysis of the results, Coulson and Silberman reported that:

1. There was no significant difference between the control and experimental groups in the pre-test.

2. A comparison of mean scores and standard deviation for the experimental and control subjects on the post-test showed that the experimental group was superior to the control group. The difference found was significant at the .01 level.

3. The second administration of the post-test criterion to the experimental subjects (as a measure of retention) showed no significant difference on the total criterion test (multiple choice plus constructed response).

4. The multiple-choice response made took significantly less time than the constructed response made. However, no significant difference was obtained between response modes on the criterion test.

5. Small steps required significantly more time but also yielded significantly higher test scores than large steps on the constructed-response criterion sub-test.

6. The branching conditions required less training than non-branching but were not significantly different on the criterion test.

It should be pointed out that while small steps yielded better results than the large, this should be weighed against the significantly greater training time required by the small-step trainees than by the large-step trainees. This would suggest that the choice of item-step size for any teaching situation must be determined in part, at least, by the extent to which required

training time is a critical consideration.

Further, since the branching and skipping procedures were extremely crude and depended upon manual control, any conclusions regarding the superiority of either branching or non-branching must be viewed with caution.

While this study would seem to indicate that the experimental subjects did learn something from teaching-machine instruction, no real conclusion can be drawn as to the effectiveness of such instruction over even unguided reading in the same area, since the control groups received no instruction of any kind.

Another study whose purpose was to determine which of the two response modes (multiple choice or constructed response) is the more efficient was that carried out by Fry (7).

In this study, Spanish vocabulary was taught to 153 ninth grade, English-speaking students, using a teaching-machine device which could be programmed to handle either multiple-choice items or items requiring the construction of an answer.

Both experimental groups were selected from within the same class by simply shuffling the programs, constructed and multiple choice, before distributing them to the class. Thus it was assumed that intelligence, motivation, socio-economic and other factors would be randomized.

A pilot study of approximately 50 pupils, who worked to a criterion of mastery, showed that the constructed-response group learned better, but also took more time. Thus the author attempted to control the time variable.

Both methods, constructed response and multiple choice, were used in the same classroom at the same time with three conditions being imposed on each method group.

Under condition 1, the groups worked to a criterion of mastery. That is, each student continued working until he had made two correct responses to each of the 16 items on the program.

Under condition 2, limited working time was the control factor. Both method groups were given the same working time. Condition-2 students were a randomly selected section of the same class as the condition-1 students.

Condition 3 was aimed at controlling both time and number of repetitions. Again there were two groups—one constructed their answers and the other selected the correct response from a multiple choice of responses. The same classes were used for condition 3 as for conditions 1 and 2. This part of the study was conducted several weeks after conditions 1 and 2.

The experimental variable under all three conditions was the method of responding. That is, recognition versus construction of answers. The Spanish words and phrases used in the programs were selected from the last few chapters of the regular class text. The same words and phrases were used for conditions 1 and 2. Those for condition 3 were selected

from the unused final portion of the students' text.

All students under all conditions were evaluated on a post-test, in which half the items were multiple choice and half were constructed response. In order to measure learning, both immediately and after a delay, two forms of the post-test were constructed. One form was administered as soon as the students had completed the training, and the other two days later.

Analysis of the results indicated that:

1. As a method for teaching students to construct answers, that is, for unaided recall, the constructed-response method was significantly superior to the multiple-choice method. However, the constructed-response group used significantly more training time to reach the criterion of mastery.

2. When time was controlled, the constructed-response method was significantly superior to the multiple-choice method on the constructed-response sub-test. However, no significant differences were found on the multiple-choice sub-tests.

3. When both time and number of repetitions were controlled, the constructed-response method was significantly superior to the multiple-choice method on the constructed-response sub-tests. Again, however, there were no significant differences found on the multiple-choice sub-tests.

On the basis of these results, Fry concluded that if recall is the criterion for learning, the constructed-response method is superior to the multiple-choice method, given conditions similar to those prevailing in this experiment. He also concluded that the constructed-response program will require more training time than the multiple-choice program.

It should be noted when considering these conclusions that, while there were significant differences between the two methods on the constructed-response sub-test, there were none found on the multiple-choice sub-tests. This suggests that Fry's conclusions might only hold for those conditions where recall is measured by constructed-response tests.

Unfortunately, Fry failed to state whether the words and phrases used for the three conditions were examined to ascertain if they were all of the same difficulty level. If the words and phrases were not of the same difficulty level, this may well have influenced his experimental findings.

NEGATIVE REINFORCEMENT

Only one study has been found in the literature that explicitly sets out to investigate the effects of negative reinforcement on programmed teaching. In this study, Melaragno (14) pointed out that various authors such as Gagne and Bolles, and Skinner contend that reinforcement should be positive, that the teaching sequence should contain maximum positive reinforcement and minimum negative reinforcement. Other authors (Coulson [4]), on the other hand, contend that it is on the basis of the student's errors

that his difficulties can be located. Programmers who hold to this view would build into their design some number of failures that will be accompanied by negative reinforcement. It was to investigate the effect of such negative reinforcement that Melaragno carried out the following study.

In this investigation, 28 subjects from freshman mathematics classes of a junior college were randomly assigned to three groups.

The instructional materials were 50 multiple-choice items, prepared so as to be extremely easy to answer. The subject matter consisted of the names and English translations of five symbols commonly used in symbolic logic, and the use of these symbols in translating from symbology to English and from English to symbology.

Five ambiguous items were also prepared, designed to look like legitimate instructional items, but containing no correct answers. These were used as a vehicle for the presentation of positive or negative reinforcement during the instructional period. A criterion test covering the material taught was also constructed.

For the first two groups, the five ambiguous items were distributed among the 50 instructional items in such a way that one ambiguous item appeared between the sixth and seventh instructional items. For the third group, the five ambiguous items were placed consecutively in the middle of the 50 instructional items. The subjects in group 1 were signaled that they were correct, no matter which choice they gave for the five ambiguous items. The subjects in group 2 and group 3 were signaled that they were incorrect, no matter which choice they gave. The design, then, was such that group 1 received 55 instances of positive reinforcement; group 2, 50 instances of positive reinforcement and five spaced instances of negative reinforcement, and group 3, 50 instances of positive reinforcement and five massed instances of negative reinforcement. All answers to the instructional items were signaled as correct, even if the answers given were incorrect. In this way, all subjects received positive reinforcement on all instructional items. The criterion test was given immediately upon completion of the 55 items.

A comparison of the results of the three groups indicated that:

1. There was no significant difference between the group receiving all positive reinforcement and the group receiving spaced negative reinforcement.

2. The scores of the groups receiving all positive reinforcement and spaced negative reinforcement were significantly greater than those of the massed negative reinforcement group.

An inspection of the tests for group 3 (massed negative reinforcement) revealed that this group not only failed to learn the symbol in which the massed negative reinforcement was placed, but also failed to learn the preceding and succeeding symbols.

Melaragno concluded from this experiment that the presence of some spaced negative reinforcement within a teaching sequence does not hinder a student's learning, but that a number of negatively reinforcing items, appearing sequentially, will tend to depress a student's learning.

While the findings of this study are of significant interest, the small size of the sample and short duration of the teaching sequence would seem to limit the generality of Melaragno's conclusions. However, several implications seem evident. One, a program which uses errors (and negative reinforcement) as a means of identifying a student's difficulty area will not penalize the student by virtue of this negative reinforcement. Secondly, any given student should not be allowed to receive too many instances of negative reinforcement, one after the other. This, of course, means that the decision to show the student easier instructional items should be made before the number of the student's consecutive errors has become very large.

EFFECTS OF GRADE LEVEL

The final study to be reviewed, while not completed at the time of this writing, has brought to light several interesting findings. The study referred to is the Automated Teaching Project (2), an extensive investigation, the objectives of which are to create, test and make available for experimental purposes, automated teaching programs for use in elementary and secondary schools. This project is under the direction of P. K. Komoski, of the Collegiate School, and B. F. Skinner.

The research reviewed here was designed to investigate the effect of grade level on learning by means of automated teaching methods. A total of 74 male subjects from the Collegiate School in New York City (25 ninth-grade students, 23 tenth-grade students, and 26 eleventh-grade students) were used. The mean Otis I.Q. for the entire group of subjects was 126.8.

The machines used were of the Skinner-type write in. Units of mathematics were programmed in sections of 150 frames per section.

Each subject was individually assigned a schedule that permitted him to work on the project for two 40-minute school periods daily.

On the day following completion of the program, each student took two tests: a post-test on the material covered in the program and a "transfer" test, comprised of items which were not specifically covered, but which demanded the concepts and skills developed by the program for successful completion. Thus, three measures were available for each subject: a post-test score, a measure of the amount learned; a transfer-test score, a measure of the ability to transfer the materials after having proceeded through the program; and the amount of time the subject took to complete the program.

An analysis of the results revealed the following:

1. No significant effect of grade level upon learning was found.

2. A significant effect of grade level upon the subject's ability to transfer the materials was found.

3. No significant effect of grade level upon the length of time needed to learn was found.

4. Observations based on the frequency distributions indicated that the program was ineffective with certain students.

On the basis of these results, the authors concluded that the traditional concepts that certain materials are best taught at a given grade level need serious re-examination. They further concluded that since grade level is positively correlated with transfer of learning, the most suitable grade level for teaching a particular unit of subject matter should be based on the student's transfer ability and not on a "vague notion of difficulty of the material."

While the foregoing study does indeed have implications for programming of material, it should be noted that the subjects used in the experiment included no low-ability students. This restriction would seem to limit any conclusions drawn from the study to the ability range represented in the sample of subjects. Also, since it can not be assumed that the subjects in this experiment were a random selection of school children, it would be difficult to generalize findings based on this sample to a larger, more universal population of school children. Further, since the program was limited to mathematics, one could well wonder if the same results would be found in other curriculum areas. The small sample and short duration of the program should also be taken into consideration.

Summary and Conclusions

On the basis of the few studies conducted to date on the effectiveness of teaching machines in the classroom, certain suggestions and reservations seem evident.

1. While two studies, Keislar (11) and The Automated Teaching Project (2), indicated that automated teaching devices were effective with brighter students, only one study, Porter (16), seemed to indicate that such devices could be used effectively with all ability ranges.

2. Two studies, Porter (16) and The Automated Teaching Project (2), showed that not all students were equally able to handle the mechanical aspects of the devices.

3. One study, Automated Teaching Project (2), suggested that confirmation of correct responses may not automatically be a reinforcer, as certain authors have supposed it to be. This raises an interesting point. Theoretically, learning can not take place without reinforcement. The most commonly used reinforcement in an automated teaching program is confirmation of

each answer made by a student. Such confirmation is a secondary, rather than a primary, reinforcement. That is, confirmation is not reinforcing in itself, but must be previously conditioned as a reinforcer. If the necessary conditioning has not taken place, the learning dependent on this conditioned reinforcer will be greatly reduced. Thus, it would seem that one problem of automated teaching may be to maintain the strength of the secondary reinforcers used. It may well be that one of the major tasks of the classroom teacher will be to condition the proper secondary reinforcers.

4. While one study, Fry (7), indicated that the constructed-response mode of programming was the most effective, another study, Coulson and Silberman (5), suggested there was little difference between this mode and the multiple-choice response method.

5. There was evidence (Porter [16], Amsel [1], Melaragno [14]) suggesting that it may not always be necessary or even desirable to minimize errors in a teaching-machine program and that the use of partial or continuous reinforcement schedules should be decided by the nature of the activity the training is to promote.

6. Since all of the studies to date have used only relatively small samples of subjects, one may well speculate as to the generality of their findings and conclusions. Certainly, since few of the studies have attempted to compare the teaching-machine method of instruction with more conventional methods, little can be concluded as to the *relative* effectiveness of automated teaching devices.

What, then, have these studies revealed? There may be some question as to the general applicability of these devices, since there is some evidence that all students could not use them equally well. Further research is necessary before one may assume that these devices will prove equally effective with all students. There is some evidence to indicate that no one type of programming, insofar as type of response is concerned, is better than another. Rather, the evidence seems to suggest that the type of program will have to depend on the aim of the particular instruction. There is evidence, also, that the presence of some negative reinforcement is not detrimental to learning, but rather that it may well be necessary in order to find the student's areas of weakness. An interesting point that certainly deserves further study is the evidence indicating that there may be need to reconsider the present concept of grade level in programming. Unless the necessity for preliminary learning in a particular subject matter can be demonstrated, the grade at which learning occurs may be irrelevant. Further, there is evidence to suggest that determining the grade level at which to teach a particular unit of subject matter should be based on the ability of the students to transfer the material which has been mastered, and not on a vague notion of difficulty of the material.

REFERENCES

1. Amsel, A. "Error Responses and Reinforcement Schedules in Self-Instructional Devices," *Teaching Machines and Programmed Learning* (edited by A. A. Lumsdaine and R. Glaser). Washington: National Education Association, 1960; pp. 506-516.
2. Automated Teaching Project; Research Summary Number 1. Collegiate School, New York, N.Y.
3. Brown, G. I. and Hodgkinson, H. L. A note concerning "An Application of Recent Developments in Psychology to the Teaching of German" by C. B. Ferster and S. M. Sapon. *Harvard Educational Review*, 28:156-157, 1958.
4. Coulson, J. E. *An Experimental Teaching Machine for Research at SDC. TM-415*, October 1959. System Development Corporation, Santa Monica, California.
5. Coulson, J. E. and Silberman, H.F. "Effects of Three Variables in a Teaching Machine," *Journal of Educational Psychology*, 51:135-143, 1960.
6. Coulson, J.E. and Silberman, H. F. *Results of Initial Experiment in Automated Teaching*. System Development Corporation, Santa Monica, California.
7. Fry, E. B. *Teaching Machines: An Investigation of Constructed Versus Multiple Choice Methods of Response*. Unpublished doctoral dissertation, University of Southern California, January 1960.
8. Fry, E. B., Bryan, G. L. and Rigney, J. W. *Teaching Machines: An Annotated Bibliography*. Washington, D.C.: Dept. of Audio-Visual Instruction, 1960.
9. Ferster, C. B. and Sapon, S. M. "An Application of Recent Developments in Psychology to the Teaching of German," *Harvard Educational Review*, 28:58-69, 1958.
10. Hively, W. "Implications for the Classroom of B. F. Skinner's Analysis of Behaviour," *Harvard Educational Review*, 29:37-42, 1959.
11. Keislar, E. R. "The Development of Understanding in Arithmetic by a Teaching Machine," *Journal of Educational Psychology*, 50:247-253, 1959.
12. Little, J. K. "Results of Use of Machines for Testing and for Drill, Upon Learning in Educational Psychology," *Journal of Experimental Education*, 3:45-49, 1934.
13. Lumsdaine, A. A. and Glaser, R. (editors). *Teaching Machines and Programmed Learning*. Washington: National Education Association, 1960.
14. Melaragno, R. J. *Results of Experiment on Negative Reinforcement in Automated Teaching*. April 1960. System Development Corporation, Santa Monica, California.
15. Porter, D. "A Critical Review of a Portion of the Literature on Teaching Devices," *Harvard Educational Review*, 27:126-147, 1957.
16. Porter, D. "Some Effects of Year Long Teaching Machine Instruction," *Automated Teaching: The State of the Art*. (Edited by E. Galanter) New York: John Wiley and Sons, 1959.
17. Pressey, S. L. "A Simple Apparatus Which Gives Tests and Scores and Teaches," *School and Society*, 23:373-376, 1926.
18. Pressey, S. L. "A Machine for Automatic Teaching of Drill Material," *School and Society*, 25:549-552, 1927.
19. Pressey, S. L. "A Third and Fourth Contribution Toward the Coming 'Industrial Revolution' in Education," *School and Society*, 36:668-672, 1932.
20. Skinner, B. F. "The Science of Learning and the Art of Teaching," *Harvard Educational Review*, 24:86-87, 1954.
21. Skinner, B. F. "Teaching Machines," *Science*, 128:969-977, 1958.
22. Stolurow, L. M. "Teaching Machines and Special Education," *Educational and Psychological Measurements*, 20:429-448, 1960.

The Diagnostic Use of WISC Subtest Patterns

KENNETH D. HOPKINS AND WILLIAM B. MICHAEL

The users of the Wechsler measures of intellectual function disagree as to how efficient the various scales are as diagnostic aids. Research studies reported on this problem have revealed conflicting and, for the most part, inconclusive results. Klebanoff (4), Parker (5), and Wechsler (7) have maintained that diagnostic significance exists for certain of the various subtest combinations. Beck (1) and Frank (2) have reported negative results with subtest pattern analysis.

Since the WISC is at present the test most widely used by school psychologists and psychometrists, the importance of knowing the reliability of subtest difference scores is apparent. Any survivor of an elementary testing course knows that if such differences are lacking in reliability, they cannot be valid. Based on the data furnished by Wechsler (5), estimates of the reliabilities of the various subtest difference scores are presented in Table I. In terms of the statistical models furnished by Kelley (3), the observed magnitude of differences between scores in any two subtests required, in the individual case, to be significant at the .05 and .01 significance levels are cited. For example, for a difference to be considered significant (.01 level) between the Vocabulary and the Picture Completion subtests, a scale-score difference of five points must be obtained.

In Table I the difference scores are shown to have low reliabilities. Yet, it is the difference score that is purported to yield diagnostic evidence. Without reliable difference scores confidence cannot be placed in the subtest differences. Reliable difference scores are associated only with highly reliable tests that are intercorrelated to a relatively low degree. Since such conditions are not found for the Wechsler scales, one must then obtain very large score differences before rejecting the hypothesis of chance variation.

The advocates of diagnostic subtest usage have frequently been inordinately vulnerable to commission of a Type I error. Some have generalized findings from group comparisons to the individual case, without giving adequate consideration to the great increase in difference required for significance with a sample of one. Wechsler (7) presents examples of subtest profiles purported to be characteristic of various clinical groups. In several

Kenneth D. Hopkins is assistant professor of educational psychology at the University of Southern California, and also serves as research consultant for the instructional television project of the Anaheim City School District. Dr. Hopkins, who formerly served as school psychologist with the Hudson School District, obtained his Ph.D. degree in 1961 from the University of Southern California.

William B. Michael has been associated with the University of Southern California since 1952, and is currently professor of psychology and education as well as director of the test bureau. Dr. Michael, who formerly taught at Princeton University and San Jose State College, obtained his Ph.D. degree in 1947 from the University of Southern California.

TABLE I

Reliability of Difference Scores Between Subtests (Above Diagonal) and Scale Score Difference Required for Significance (.05 and .01 Significance, Below Diagonal)[1]

	Inf.	Comp.	Ari.	Sim.	Voc.	D. S.	P. C.	P. A.	B. D.	O. A.	Cod.[2]	Ver.	Perf.	F. S.
Information		.33	.42	.41	.42	.51	.54	.50	.68	.60	.34	.30	.71	.55
Comprehension	4.1 7.6		.57	.38	.22	.42	.52	.46	.64	.51	.48	.48	.45	.48
Arithmetic	3.5 4.6	3.9 5.0		.56	.67	.48	.63	.57	.72	.61	.52	.67	.66	.66
Similarities	3.6 4.7	4.0 5.2	3.5 4.5		.64	.51	.60	.59	.74	.63	.56	.59	.71	.66
Vocabulary	3.2 4.1	3.5 4.6	2.9 3.8	3.1 4.1		.52	.59	.57	.80	.61	.55	.64	.69	.83
Digit Span	4.6 6.0	4.8 6.1	4.4 5.7	4.5 5.9	4.2 5.4		.58	.48	.59	.40	.39	.55	.56	.54
Picture Completion	4.3 5.6	4.6 6.0	4.2 5.4	4.3 5.6	3.9 5.0	5.1 6.6		.52	.56	.43	.51	.65	.57	.60
Picture Arrangement	4.1 5.3	4.4 5.7	3.9 5.1	4.1 5.3	3.6 4.7	4.9 6.4	4.6 6.0		.57	.53	.43	.60	.57	.55
Block Design	3.4 4.4	3.7 4.8	3.2 4.1	3.4 4.4	2.8 3.6	4.3 5.6	4.1 5.3	3.8 5.0		.39	.59	.81	.65	.75
Object Assembly	4.4 5.7	4.7 6.1	4.3 5.6	4.4 5.7	4.0 5.2	5.2 6.7	4.9 6.4	4.8 6.2	4.2 5.2		.47	.67	.50	.60
Coding	4.7 6.1	4.9 6.4	4.5 5.9	4.6 6.0	4.3 5.6	5.4 7.0	5.2 6.7	5.0 6.5	4.4 5.7	5.3 6.9		.59	.58	.57
Verbal	2.9 3.7	3.8 4.9	2.6 3.4	2.8 3.7	2.1 2.8	3.9 5.1	3.6 4.7	3.4 4.4	2.4 3.1	3.8 4.9	4.1 5.3		.77	.36
Performance	3.4 4.3	4.1 5.3	3.1 4.0	3.2 4.2	2.6 3.4	4.2 5.5	3.9 5.1	3.7 4.8	2.9 3.7	4.1 5.3	4.4 5.6	11.5* 16.0		.20
Full Scale	2.9 3.8	3.8 5.0	2.7 3.5	2.9 3.7	2.2 2.8	4.0 5.2	3.6 4.7	3.4 4.4	2.5 3.2	3.8 5.0	4.1 5.3	9.5* 11.5	11.7* 15.5	

[1]Computed from data given by Wechsler (6); 100 boys and 100 girls at age 10 and one-half.
[2]Since no data were available at 10 and one-half age level for coding, data for 8 and one-half age level were substituted.
*Difference given in terms of IQ scale scores.

of the "characteristic profiles" (e.g., organic brain disease, anxiety states) there are no subtest differences large enough to reject the null hypothesis of random variation at the .05 level. Obviously, before offering diagnostic interpretations, one should be reasonably sure that there is something other than chance variation to interpret.

Although it is realized that the degree of reliability of any of the subscales will vary in different samples and under various conditions of testing, it is hoped that the entries in Table I will serve as an approximate guide to determining whether the size of obtained difference scores is sufficient

(Continued on page 130)

Some Participant Reactions To A Cooperative Approach to the Study Of Educational Problems

CASEEL D. BURKE

In these days of apparent popularity of the cooperative approach to the study of school problems, there may be value in investigating the benefits to be gained from this practice of school people and laymen getting together to identify and work on matters of concern to the schools. The usual measure of success in these undertakings is the nature or quality of the recommendations, proposals, or resolutions developed by the participants. Beyond this type of gain, however, is the possibility that substantial benefits may come from analyzing the *process* of the cooperative effort. Knowing how the participants react to their cooperative experience may help educators, citizens committees, and others in their planning for joint ventures in solving educational problems. This study deals with the values and inner-workings of the cooperative approach as sensed by the participants in the process.

Do participants in an education study group consider the experience worthwhile? Are they willing to repeat the experience? Is there freedom of discussion? Do certain people dominate the group? What motivates people to join an educational study group? Some answers to these and other similar questions are a by-product of the Cooperative Study of Elementary Education in California, sponsored by the California Elementary School Administrators Association (CESAA) and participated in at one stage or another by a total of more than 8,000 citizens of that State.

Caseel D. Burke has been with the department of education at Utah State University for twelve years, and is currently chairman of the division of teacher education. Dr. Burke has taught in both elementary and high schools, and also served as the assistant executive secretary of the California Elementary School Administrators Association. While on leave from Utah State, he was consultant in education and part-time lecturer in education at the University of California, Berkeley, and also served as advisor in elementary teacher education, National Teachers College, University of Tehran, Iran. Dr. Burke obtained his Ed.D. degree from the University of California, Berkeley, in 1953.

During the "local study group phase" of the Cooperative Study there were organized throughout the State of California nearly 400 groups which averaged about 13 members each and which had an approximate membership ratio of five professional educators to eight persons of other occupations. Their problem was to determine the "characteristics of the good elementary school"[1]—the kind they would like to have and would be willing to support. The groups averaged between three and four meetings each in discussing the problem. In their deliberations they reacted to a tentative outline previously developed through extensive effort by educators and lay people. A questionnaire soliciting their reactions to the study group experience was sent to each of the approximately 4,500 study group participants and responses were received from 1,463. The following is an analysis of their reactions:

The Values

What is to be gained by participation in an education study group? In answer to this question 61 per cent of the participants in occupations other than education indicated they had gained a better understanding and appreciation of problems facing the schools. The professional educators gained less in this respect, as is indicated by their approximately 38 per cent response of a similar nature.

About one in twenty (4.4 per cent) of the respondents felt the experience was mostly a waste of time. The men were somewhat more of this opinion than were the women, and those in other occupations held the opinion to a greater degree than did the educators.

"I would be willing again to join a study group with a similar purpose" was the attitude of approximately 60 per cent of the respondents. The educators were a bit less enthusiastic than were those in other occupations. The men, also, were somewhat less willing to participate again than were the women.

Nearly three-fourths of the respondents expressed the belief that a cooperative approach is a good way to attempt to solve school problems.

Freedom versus Domination

The question *"Did you feel completely free to express your ideas?"* brought an affirmative response from more than 92 per cent of the study group members. Only slightly more than 1 per cent felt they did not have

---

[1]CESAA. *The Characteristics of the Good Elementary School.* Burlingame: The California Elementary School Administrators' Assn. 1957. 50 pp.

TABLE I

The Values Gained from the Study Group

Response[1]	*Total Respondents (1463)*	*Males (456)*	*Females (1007)*	*Educators (500)*	*Other Occupations (963)*
It helped me to better understand and appreciate the problems facing the schools	52.9%	45.2%	56.4%	37.6%	61.0%
It helped most group members to better understand and appreciate the problems facing the schools	58.8	59.2	58.6	62.6	52.8
It was mostly a waste of time	4.4	5.0	4.2	2.8	5.3
I would be willing again to join a study group with a similar purpose	59.4	51.3	63.1	52.2	62.5
I believe a cooperative approach is a good way to attempt to solve problems concerning the schools	73.8	72.1	74.5	79.0	77.5

[1]Since participants were asked to respond to as many items as they desired, the percentages total more than 100% in most cases. This is true also of Tables II, III, and IV.

the opportunity to express themselves fully.

Nearly one-half of the respondents asserted that no one attempted to dominate or influence the discussion; whereas nearly 8 per cent felt that some persons had exerted special influence. Approximately 3 per cent noted that certain persons did dominate and unduly influence the discussions and decisions.

Apparently about 40 per cent of the respondents, in not choosing to answer either way on the matter of domination, sensed no particular problem here. The educators seemed to be even less concerned in this respect than were the others, since over 50 per cent of them did not respond on the matter of domination.

The fact that about 10 per cent of the participants in the discussions noted some domination or unusual influence exerted by certain members of the group likely would not seem out of the ordinary to anyone who has participated in group discussions of various kinds. Considering the domi-

TABLE II

Freedom and Domination in Study Group Discussions

Response	*Total Respondents (1463)*	*Males (456)*	*Females (1007)*	*Educators (500)*	*Other Occupations (963)*
I felt completely free to express my ideas	92.2%	91.2%	92.7%	91.2%	93.0%
I did not have as much opportunity to express myself as I would have liked	1.2	1.1	1.3	2.0	0.8
No one attempted to dominate or influence the discussions	47.3	38.2	51.4	38.4	52.0
Some persons had special influence in the discussions and decisions of the group	7.5	7.7	7.4	6.8	7.9
Certain persons dominated the discussions and unduly influenced the decisions of the group	2.9	3.1	2.9	2.2	3.3

nance in numbers of educators over any other single occupation group, some might expect to find them imposing their views on the others. The evidence, however, tends to discount the presence of dominance in the discussions from whatever source it might have come.

Sources of Invitation

In an effort to gain as extensive participation as possible in the study group phase, the sponsors of the Cooperative Study extended invitations as widely throughout the State as seemed practicable. Key people in the study were asked to submit prepared releases through local newspapers in all geographical areas. Also, the major civic, professional, and educational organizations, including school districts and the Parent-Teacher Associations, were asked to enlist the support of their local units in encouraging wide participation.

Contacts made by persons representing the schools proved most fruitful in enlisting group members, as 63 per cent received their invitations from this source. Twenty-nine per cent received their invitations through an

TABLE III

Sources of Invitation to Participate

Response	*Total Respondents (1463)*	*Males (456)*	*Females (1007)*	*Educators (500)*	*Other Occupations (963)*
I was invited by a friend	4.2%	3.7%	4.4%	2.4%	5.4%
I was invited through an organization of which I am a member	29.3	26.8	30.4	27.2	30.0
I was invited by someone representing the schools	63.4	66.9	61.9	64.4	62.2
I was invited through an informal notice in the press or otherwise	3.0	2.5	3.2	0.8	4.2

organization to which they belonged; whereas informal invitations in the press and on the radio accounted for only 3 per cent of the participants. Four per cent came on the invitation of a friend.

It is obvious that equal effort was not expended in each of these methods of contact, and comparisons are thus meaningless. However, there is evidence to show that because of many factors, such as convenience of contact, size of the organization, nature and purpose of the organization, and the concern of individuals involved, some methods of recruitment were superior to others.

Reasons for Participating

Since all action is based on motivation of one kind or another, there should be value in knowing why people respond when the call comes to consider school problems.

To the question *"Why did you participate in the study group?"* 70 per cent answered they considered this a chance to contribute something to the schools; 12 per cent felt it was their obligation; and 37 per cent stated they were curious. Curiosity was a much greater motivating factor for the women than for the men; and the educators, many of whom were women, were considerably less curious than were those from other occupations.

Nearly 3 per cent of the respondents participated in the study groups because they felt they couldn't decline gracefully. This was slightly more true of educators than of those from other occupations.

TABLE IV

Reasons for Participating

Response	*Total Respondents (1463)*	*Males (456)*	*Females (1007)*	*Educators (500)*	*Other Occupations (963)*
I considered this a chance to contribute something to the schools	69.5%	70.2%	69.2%	67.0%	70.8%
I considered it an obligation when asked to join	12.4	12.1	12.5	13.4	11.8
I was curious	37.0	26.1	41.9	24.6	43.3
I couldn't decline gracefully	2.8	2.2	3.0	3.0	2.6

Summary

Of the approximately 4,500 persons who participated in the local study group phase of the Cooperative Study of Elementary Education in California, 1,463 responded to a number of questions concerning what they thought of the study group experience. An analysis of their responses reveals the following information:

1. A substantial majority of the participants felt they had gained a better understanding and appreciation of school problems.

2. Nearly three-fourths of the participants felt the cooperative approach is a good way to attack school problems.

3. A substantial majority expressed a willingness to engage again in an activity of this kind.

4. Nearly all of the participants felt they had been free to express their own ideas to the group.

5. About 10 per cent of the respondents sensed some domination of the discussions by certain members.

6. Best results in enlisting people to participate in the study groups came through the facilities of established organizations—the schools, parent-teacher associations, and civic welfare groups.

7. More than two-thirds of the participants joined study groups from the desire to contribute something to the schools. Curiosity as to what the activity was about was a motivating force for a considerable number.

A Comparative Study of Children's Play In America and Japan

MAY V. SEAGOE AND KEN MURAKAMI

The universality of children's play and its significance in the transmission of the culture is generally recognized (3). Specific activities differ in various cultures, yet the inevitable emergence of play and certain basic forms it takes suggest origins which are deeper than the culture in which the child finds himself. Theories of play (7) include release of surplus energy, recreation or restoration of sense of well-being after fatigue, practicing instincts in the process of growing up, and recapitulation of the experience of the race. Whatever the origin of play, it contributes to physical development, emotional release, socialization, learning, and inventing (4).

Studies of play, however, have been largely descriptive and highly subjective in nature, tracing patterns in a single culture. Cross-cultural controlled research using relatively sophisticated modern cultures rather than primitive groups is conspicuously absent. The common factors which appear in different cultures, the reflection of cultural differences which children's play provides, and ways in which data on play from different cultures may be objectively analyzed are the concern of this paper.

A comparative study (8) of children's play in America and Japan, undertaken in 1960, is described in this article.

Research Plan

"Play" was equated to *Asobi* for the Japanese group. Though *Asobi* is somewhat broader in scope than "play," the essential "fun" element is stressed in both (3). The hypothesis tested was that the factors of age, sex, and type of community had different impact on children's activities in the two cultures.

Subjects were chosen to represent comparable samples from Japanese and American cultures. Age (first and sixth grades), sex, type of community

---

Ken Murakami has been in business, in addition to teaching English, in Takarazuka, Japan, for the past year. He received his M.A. degree in 1960 from the University of California, Los Angeles, and this article is taken from his thesis.

May V. Seagoe has taught at UCLA since 1934 and has also served as assistant dean since 1958. Dr. Seagoe, whose articles have been widely published, obtained her Ph.D. degree from Stanford University in 1934.

(rural and urban), and season (date of sample), were held constant. In addition, a similar sample from an American school in Japan was studied. Using 15 boys and 15 girls from two American schools (rural and urban), the same number from two comparable Japanese schools, and the group from the American School in Japan, 300 cases were selected. Validity of the comparison rests not only on objective data but also on the personal knowledge of the Japanese schools on the part of the author (Murakami) and a systematic search for similar communities in California.

Data were collected in controlled interviews with the subjects. Questionnaires were dismissed as a method of data collection because of lack of control; checklists because they demand too much reading facility for young children; and direct observation because of its demands in time and place. Interviews in America were conducted by Murakami, those in Japan by three teachers under his detailed instruction.

The interview was developed through several preliminary trials with children. The data used consisted of three choices each to four questions: **What do you spend most of your time playing at school? At home? What do you like to play most? Least?**

Specific names of play activities from the questionnaires were then classified into developmental categories similar to those derived by Renfoe (9). Other types of categorization, such as play material or bodily involvement, were rejected in favor of one stressing the socialization process. One of Renfoe's categories (ritual games) was broken into three to reflect more clearly an important cultural difference. The categories finally used were: dramatic play, "circle" games, group activities, individual activities, team competitive games, individual competitive games, games of intellect, collecting, and passive participation.

Data were analyzed for each of the five samples in terms of age differences, sex differences, and type of community. The per cent of answers in each category was computed, and the "t" test applied to differences between groups for each category.

Results

Table I shows percentages of response to the four questions summarized by age. Level of confidence for significant differences between Japanese and Americans of the same age are indicated by *a* or *b*, significant difference for first and sixth grade children within the *same* culture by *c* and *d*.

Group activities are clearly more characteristic of Japanese children of both ages, and team competitive games a little less distinctly characteristic of Americans of both ages. Age differences emerge in both groups in the movement from individual activities in first grade to team competitive games in sixth grade, though the change is more marked for the Americans. There is also a unique American emphasis in first grade on dramatic play and

TABLE I

Age Differences in Play

Play Category	*Question*	*Japanese*		*Americans*		*Americans in Japan*	
		1	*6*	*1*	*6*	*1*	*6*
Dramatic play	School	0	0	4	0	17	0
	Home	21	5	24	8	44	0
	Liked	22	11	17[c]	2[c]	28	0
	Disliked	14	5	15	20	29	21
"Circle" games	School	0[a]	0	13[ac]	0[c]	2	1
	Home	0	0	6	1	1	0
	Liked	0	0	8	3	0	0
	Disliked	0	0	7	11	0	0
Group activities	School	63[b]	46	23[b]	32	53	16
	Home	48[b]	44	13[b]	23	11	17
	Liked	50[b]	36	18[b]	17	27	9
	Disliked	51[b]	59	19[b]	22	21	21
Individual activities	School	27[d]	4[d]	36[d]	2[d]	26	8
	Home	18[ac]	4[c]	41[ac]	15[c]	30	9
	Liked	13	4	27	11	22	19
	Disliked	17	7	17	8	21	0
Team competitive games	School	3[ad]	29[bd]	22[ad]	64[bd]	0	67
	Home	2[c]	17[ac]	13[c]	45[ac]	2	33
	Liked	5[d]	29[ad]	17[d]	60[ad]	5	42
	Disliked	8[a]	13	29[a]	30	2	33
Individual competitive games	School	6	13	1	2	0	8
	Home	4	10	1	2	2	15
	Liked	3	11	4	5	3	15
	Disliked	7	12	4	8	14	2
Games of intellect	School	0	0	0	0	2	0
	Home	0	1	1	3	6	25
	Liked	0	0	4	1	8	10
	Disliked	0	0	7	1	4	21
Collecting	School	1	1	0	0	0	0
	Home	0	2	0	0	0	0
	Liked	0	0	0	0	0	0
	Disliked	0	0	0	0	0	0
Passive participation	School	0	7	1	0	0	0
	Home	7	17	1	3	4	1
	Liked	7	9	5	1	7	5
	Disliked	3	4	2	0	9	2

[a] Difference between *cultures* for same age significant at the .05 level.
[b] Difference between *cultures* for same age significant at the .01 level.
[c] Difference between *ages* for same culture significant at the .05 level.
[d] Difference between *ages* for same culture significant at the .01 level.

"circle" games, one that disappears by the sixth grade.

Table II presents similar data differentiated by sex. Here again the tendency for both boys and girls in Japan to engage more often in group activities is clear. American boys exceed Japanese in their preference for team competitive games. This tendency is not equally shared by American girls who fail to carry over team sports to the home and more often dislike

TABLE II

Sex Differences in Play

Play Category	*Question*	*Japanese*		*Americans*		*Americans in Japan*	
		B	*G*	*B*	*G*	*B*	*G*
Dramatic play	School	0	0	2	2	9	7
	Home	11	15	8	23	17	28
	Liked	14	18	7	12	7	21
	Disliked	9	9	20	5	20	29
"Circle" games	School	0	0	3	10	0	4
	Home	0	0	1	5	1	0
	Liked	0	0	2	7	0	0
	Disliked	0[a]	0	14[a]	8	0	0
Group activities	School	48[a]	61[a]	27[a]	28[a]	31	38
	Home	41[a]	51[a]	18[a]	19[a]	9	19
	Liked	39[b]	48[b]	14[b]	23[b]	19	17
	Disliked	56[b]	55[b]	22[b]	21[b]	20	22
Individual activities	School	12	19	18	20	10	24
	Home	10	12	27	27	25	15
	Liked	6	11	21	17	22	19
	Disliked	7	18	12	13	14	9
Team competitive games	School	23[a]	9[a]	50[a]	37[a]	40	27
	Home	12[a]	7	40[ac]	18[c]	18	17
	Liked	21[b]	13	48[b]	29	29	18
	Disliked	11	9[b]	24	40[b]	18	16
Individual competitive games	School	13[a]	6	0[a]	2	8	0
	Home	12	2	2	2	13	3
	Liked	12	2	5	4	10	9
	Disliked	14	6	6	8	7	8
Games of intellect	School	0	0	0	0	2	0
	Home	1	1	2	3	16	15
	Liked	0	0	1	3	6	11
	Disliked	0	0	2	4	16	11
Collecting	School	2	0	0	0	0	0
	Home	1	1	0	0	0	0
	Liked	0	0	0	0	0	0
	Disliked	0	0	0	0	0	0
Passive participation	School	2	5	0	1	0	0
	Home	12	11	2	3	1	3
	Liked	8	8	2	5	7	5
	Disliked	3	3	0	1	5	5

[a] Difference between *cultures* for same sex significant at the .05 level.
[b] Difference between *cultures* for same sex significant at the .01 level.
[c] Difference between *sexes* for same culture significant at the .05 level.
[d] Difference between *sexes* for same culture significant at the .01 level.

them. They are intermediate between American and Japanese boys in team sports, but clearly differ from the Japanese girls in emphasis on group activities, more often using their time in individual activities or "circle" games. The stress on individual competitive sports among the Japanese boys, and the dislike of American boys for "circle" games constitute unique differences.

TABLE III

Rural-Urban Differences in Play

Play Category	*Question*	*Japanese*		*American*	
		Urban	*Rural*	*Urban*	*Rural*
Dramatic play	School	0	0	1	3
	Home	12	13	20	12
	Liked	15	18	12	7
	Disliked	9	10	14	12
"Circle" games	School	0	0	12	1
	Home	0	0	6	0
	Liked	0	0	10	1
	Disliked	0[a]	0	17[ac]	3[c]
Group activities	School	48[a]	59[a]	25[a]	30[a]
	Home	31[c]	61[bc]	19	18[b]
	Liked	33	54[b]	19	16[b]
	Disliked	43[ac]	73[bc]	16[a]	25[b]
Individual activities	School	21	11	14	24
	Home	14	8[a]	22	33[a]
	Liked	11	6[a]	12	27[a]
	Disliked	18	5	11	12
Team competitive games	School	15[b]	17[a]	46[b]	40[a]
	Home	10	10[a]	24	33[a]
	Liked	14[a]	20	39[a]	38
	Disliked	12	8[a]	28	32[a]
Individual competitive games	School	11	9	2	1
	Home	10	3	2	1
	Liked	11	2	7	2
	Disliked	14	4	7	14
Games of Intellect	School	0	0	0	0
	Home	1	1	3	2
	Liked	0	0	0	4
	Disliked	0	0	6	1
Collecting	School	0	2	0	0
	Home	0	2	0	0
	Liked	0	0	0	0
	Disliked	0	0	0	0
Passive participation	School	5	2	0	1
	Home	22[ad]	2[d]	4[a]	1
	Liked	16[ac]	0[c]	1[a]	5
	Disliked	4	0	1	1

[a] Difference between *cultures* for same type of community significant at .05 level.
[b] Difference between *cultures* for same type of community significant at .01 level.
[c] Difference between *type of community* for same type of culture significant at the .05 level.
[d] Difference between *type of community* for same type of culture significant at the .01 level.

Table III analyzes the data for rural and urban groups. Again the emphasis of both rural and urban groups in Japan on group activities is marked, though it is greater in rural than urban Japan. Similarly, the emphasis of both urban and rural groups in America on team sports is clear, though the difference between rural Americans and rural Japanese

is greater than that between urban groups. The rural American group also exceeds the Japanese in individual activities. The results on passive participation reflect largely the advent of television in the Japanese city in contrast to the country. In general, the difference between the Japanese rural and urban samples appears to be greater than between the American rural and urban subjects.

Americans in Japan generally occupy a position between Japanese and Americans, suggesting modification of cultural pressures there. More significantly, inspection of the basic tables shows: (1) greater variety in play among the Americans in Japan; (2) a total of 94 activities compared with 57 for the Japanese and 83 for Americans; and (3) 52 unique activities compared to 25 for the Japanese and 51 for the Americans. The basic tables also show that the Americans in Japan play the American games and unique games, but not Japanese games; of their 94 play forms only one type of informal social play is shared with the Japanese only, while 16 play forms are shared with Americans only and 25 with both Americans and Japanese. The data suggest further study of Americans in other cultures to see whether the modification of cultural pressure and the increased variability of behavior are present in other transplanted groups.

Conclusions

In summary, the two cultures show many similarities in age, sex, and rural-urban patterns. Often the intra-cultural difference exceeds the cross-cultural difference. Specific activities differ markedly, but game categories are similar.

The play of Japanese children stresses informal group activities and, to a lesser degree, individual competitive sports. Rural-urban differences are more marked as well.

The play of American children stresses team sports, with much greater age differentiation in activities. There is also more emphasis, particularly among young children and rural children, on informal individual activities.

The play of children in the American school in Japan reflects to a lesser degree the stress of both cultures. It also shows greater variety of activities, more unique play forms, and a greater tendency to share American games rather than Japanese.

SELECTED BIBLIOGRAPHY

1. Appleton, Lilla Estelle. *A Comparative Study of the Play Activities of Adult Savages and Civilized Children.* Chicago: University of Chicago Press, 1910.
2. Hartley, Ruth Edith. *Understanding Children's Play.* New York: Columbia University Press, 1952.
3. Hiuzinga, Johan. *Homo Ludens.* New York: Roy Publishers, 1950.
4. Kraus, Richard G., *Play Activities for Boys and Girls.* New York: McGraw-

Hill, 1957.

5. Lehman, Harvey Christian and Witty, Paul S. *The Psychology of Play Activities.* New York: Barnes and Company, 1927.
6. Mead, Margaret and Wolenstein, Martha. *Childhood in Contemporary Culture.* Chicago: University of Chicago Press, 1955.
7. Mitchell, Elmer Dayton and Mason, Bernard S. *The Theory of Play.* New York: A. S. Barnes, 1939.
8. Murakami, Ken. *A Comparative Study of Children's Play Activities in America and Japan.* Unpublished Master's thesis, University of California, Los Angeles, 1960.
9. Renfoe, Olive. *A Study of the Developmental Sequence as Revealed in the Anecdotal Records of Teachers.* Unpublished Ed.D. dissertation, University of Maryland, 1952.
10. Schwender, Norma. *Game Preferences of 10,000 Fourth Grade Children.* Unpublished Ph.D. thesis, New York, Columbia University, 1932.
11. Sullivan, Harry Stacke. *Conceptions of Modern Psychiatry.* Washington, D.C.: The William Alanson White Psychiatric Foundation, 1940, 153 pp.
12. Wang, James Daniel, *A Study of Certain Factors Associated with Children's Play Interests.* Unpublished Ph.D. thesis, Nashville, George Peabody College for Teachers, 1941.

WISC Subtest Patterns

(Continued from page 117)

for making reliable diagnostic inference. Of course, the message imparted by the entries should not serve to discourage the research worker from further exploration of hypotheses he may have obtained from differential performance, especially when he has obtained consistent differences that fall only slightly short of significance.

REFERENCES

1. Beck, H. H., and Lam, R. L. "The Use of the WISC in Predicting Organicity," *Journal of Clinical Psychology,* 11:155-158.
2. Frank, I. G. "Empirical Critique of Research with Wechsler Intelligence Scales," *Journal of Clinical Psychology,* 11:291-293.
3. Kelley, T.L. *Fundamentals of Statistics,* Cambridge: Harvard University Press, 1947.
4. Klebanoff, S. C., Singer, J. L., and Wilensky, H. "Psychological Consequences of Brain Lesions and Ablations," *Psychological Bulletin,* 57:132-155.
5. Parker, J. W. "The Validity of Some Current Tests for Organicity," *Journal of Consulting Psychology,* 15:62-67, 1957.
6. Wechsler, D. *Wechsler Intelligence Scale for Children*—Manual, New York: Psychological Corp., 1949.
7. Wechsler, D. *The Measurement and Appraisal of Adult Intelligence* (4th ed.), Baltimore: William & Wilkins, 1958.

MEET YOUR CACER -- 8

(This is the eighth of a series of biographical sketches of the members of the California Advisory Council on Educational Research.)

Happy, but sad, is the mood of this biographical sketch. The cause of this paradox is Dr. Henry I. Weitzel's announcement of his pending retirement this summer from his position of Research Director of the Pasadena City Schools after serving ten years in that capacity. We are happy for Hank, because his retirement will not mean that he will "rust" away, for on July 5 he will become the Registrar of the California Institute of Technology. His new position will be challenging, but we hope that it will not be so confining that he will neglect his stamp collection. We also desire that he and his wife, Ida, will be able to attend the legitimate theater without fear of last minute cancellations, as was so often necessitated in the past by report deadlines. They may be able to squeeze in leisurely travel, although Hank has seen many parts of the world while in service as a Commander in the United States Coast Guard. So, there is a note of happiness.

Yet, there is always a sadness attached to a retirement—especially when it is a departure of a man who has served a school district for almost forty years, as teacher, counselor, Dean, and Research Director. There will be a void in the Pasadena City Schools, for Hank has become almost a tradition, but he will not be lost to the community, the state, or the nation. He can still participate in the dozen professional organizations to which he belongs, and he can continue his numerous writings in periodicals and textbooks.

CACER will suffer the greatest loss, even though he will continue as our Associate for the next few months. The Council will miss his ready wit and inevitable jest; his sincerity, and his honesty. People who know him realize that his integrity is beyond question. But the Council will not miss the cigar he invariably smokes.

Lots of luck, Hank and Ida. CACER is happy for you, but sad that you won't be with us—except in thoughts and spirit.

High School Students' Values: Emergent Or Traditional

O. E. Thompson

We cannot think about the future of American education without first attempting some forecast of changing values in our emerging civilization (7). Values, when thought of in terms of education, are often confused with attitudes and goals. In this discussion, values are understood to be the relatively stable pre-determiners of action that are developed in each individual through prizing, cherishing, or holding something dear. In general, values govern behavior directly. When the environment is such that value-directed behavior is not appropriate, however, the individual, by a change in attitude, usually adjusts satisfactorily to the situation. By this change in attitude the individual effects a compromise between his values and expected behavior. When the environment again permits it, the individual tends to revert to value-directed behavior.

Though the home is recognized as a potent factor in molding of the personality of children, the school is gaining in prominence for its contribution to value development. Parental influence is known to decline as the child grows older. Brown, et al (1), found that the family unit begins playing a less significant role in value formation between the ages of 10 and 16 years, with social factors becoming more important. In studying 277 delinquent boys and girls from 9 to 14 years old, Crane (2) found that before the age of 12 girls tended to identify with family members of the same sex whereas boys chose heroes in sports and in history. From 12 years on, adults close to the family became the idols of boys, while girls identified with adults outside the family circle. Havighurst and associates (4) reported that the self-concept of children is heavily influenced by associations with adults who have positions of prestige: "*The inference is clear that schools, churches*

O. E. Thompson is assistant professor at the University of California, Davis, a position he has held for six and a half years. Formerly a high school teacher and instructor at Cornell University, Dr. Thompson obtained his Ph.D. degree in 1954 from Cornell. In addition to writing numerous articles for professional journals, Dr. Thompson served as a consultant for the State Department of Education in 1956-57.

and youth-group leaders influence the ideals of youth as much or more through the presence and behaviors of teachers, clergy, and youth-group leaders as through their verbal teaching." Jacobs (9), in a review of the research during the past 15 years, found that the value patterns of students ordinarily are already set at entrance to college, changing little during the college years. Thus the elementary and the secondary schools have inescapable roles in supplementing and complementing the influence of the home in value "interiorization." These schools can affect the acquisition of values in four main ways: by reinforcing desirable values; by reducing the influence of undesirable values; by promoting a child's self-respect; and by helping him develop moral insight (5). Clearly, teachers have a potentially great influence on youth, and in marginal cases in value development their influence is likely to be decisive.

This study had as its primary objective the determination of personal value patterns of high school seniors and freshmen. Comparisons were made among such groups, stratified by grade, sex, scholastic achievement, curriculum major, church attendance, and other factors. The vocational values and desires, also measured, are discussed briefly.

Procedure

The personal values of students were measured by the *Differential Values Inventory* developed and validated by Richard Prince at the University of Chicago (8). This instrument is based on the premise that "*under the necessity of reorganizing our social structure to meet the demands of a new technology and of a spacial mobility unparalleled in modern history, our inherited system of status and roles is breaking down, while a new system compatible with the actual conditions of modern life has not yet emerged*" (6). In other words, we are in a period of transition from one major culture to another that is not yet obvious in its entirety. Getzel (3), building on the works of G. A. Spindler, organized personal values into two fields—traditional and emergent. He has identified as traditional the secular values of Puritan morality, independence, work success, and orientation to the future. Parallel to each secular class of value he placed the emergent value that is believed to be its opposite in our modern culture. Thus, in the emerging society change is taking place from Puritan morality to moral relativism, from personal independence to group conformity, from work success to sociability, and from future-time orientation to present-time orientation. Using these eight categories of values, Prince (8) developed a forced-choice questionnaire containing 64 pairs of items, each consisting of a traditional- and an emergent-value statement preceded by the words "*I ought to—.*" In each case the student identified the value statement in the pair that was the more important to him personally. Example statements are:

2. (1.) I ought to do things that other people do.
(2.) I ought to do things that are out-of-the-ordinary.
64. (1.) I ought to strive to be an expert in at least one thing.
(2.) I ought to do many things quite well but not necessarily be an expert in anything.

Each student made two scores — one for the traditional-value items chosen—the other for the emergent-value items. Obviously, the total score for each student was 64.

This instrument, plus a personal-information blank and an occupational-values survey, was completed by each senior and freshman in five Central California high schools. From this population, random samples of 100 boys and 100 girls from both the senior and freshmen classes were drawn for analysis. This study duplicates in part the research of Prince and Getzel. However, additional information was obtained and the analysis of the data is different in this study.

Discussion of Results

Most of the analyses that follow use only the traditional-value scores. There is no particular reason for choosing traditional values except that they are more easily understood by most people.

In comparing the mean traditional- and emergent-value scores for all students, it is found the difference of 2.93 is significant beyond the .05 level, indicating that these observed differences are unlikely to be sampling error. The values held by the students in these five high schools were significantly more emergent than traditional. When using this value inventory with freshmen and seniors in two private, sixteen public, and four religious schools in Illinois, Prince (8) obtained respective mean traditional-value scores for seniors of 27.50, 33.50, and 38.56. The mean of 33.50 for public school seniors in that area is somewhat higher than the mean of 30.91 for California seniors. Several reasons are possible. First, the absence of private (non-parochial) schools in this area would put the otherwise private-school type of student with high emergent values in the public school, thereby possibly depressing the mean traditional-value score. Second, the very nature of the dynamic Western culture may contribute more to emergent types of values than does the Middle-Western culture. Finally, these data were collected four years after the previously mentioned study. The general trend toward emergent types of values, predicted by Getzel, may be reflected in the California sample.

When making comparisons between the traditional and emergent values of the various groupings, significant differences were found in all groups except senior boys. Senior boys were the most traditional of all the groupings, but not enough so to be significantly more traditional than emergent.

TABLE I

Comparison of Traditional and Emergent Values of Various Groupings of High School Students

Grouping	*N*	*Mean "T" Score*	*Mean "E" Score*	*t-Value*	*Significance* .05	.01
All Students	400	30.52	33.45	5.748		*
Freshmen	200	30.14	33.80	4.867		*
Seniors	200	30.91	33.09	3.177		*
Boys	200	31.00	32.94	2.614		*
Girls	200	30.04	33.96	5.773		*
Urban Residents	265	30.56	33.43	4.743		*
Rural Residents	135	30.44	33.47	3.254		*

The consistency of the mean traditional-value scores among the various groupings is noteworthy. Not even place of residence influenced the value patterns of the students. However, it should be clarified here that in this sample "rural residents" doesn't necessarily mean farmers, for only one-third of this group listed agricultural work as the fathers' occupation.

In comparing the mean traditional values of freshmen and seniors and boys and girls, differences are observed, but in no instance are these large enough to be significant at the .05 level. These data suggest that values don't change between the freshman and senior year in high school, and that there is no major difference in values between boys and girls. Individual value categories, however, show evidence of significant changes, discussed later in this report. It appears that seniors are consistently more traditional than freshmen, and that boys are generally more traditional than girls, though the differences are not statistically significant.

A question frequently asked is: **"Are the values of a student affected by his home situation?"** Here, students were asked to record with whom they lived. About one in eight came from a home situation that did not include both a father and mother. Of the 49 in the sample from atypical homes, 32 lived with the mother and the remainder were with the father or relatives, or in foster homes. Boys from atypical homes had exactly the same mean traditional-value scores as those from typical homes (see Table II). Girls from atypical homes were considerably more emergent in values than other girls or boys from comparable home situations, but not enough different for significance at the .05 level. Thus it can be inferred from these data that the home-family situation did not influence student traditional or emergent values markedly.

Several things can be inferred from the occupation of the head of the household. It can well indicate the socio-economic level of the family, the

TABLE II

Traditional Values of Students Coming from Typical and Atypical Homes

Students Living With:	*Boys*		*Girls*		*t-Value*
	N	*Mean "T" Score*	*N*	*Mean "T" Score*	
Both Parents	179	31.00	172	30.26	0.931*
Mother, Father or Relatives	21	31.00	28	28.64	1.268*

*Not significant at the .05 level.

educational achievements of fathers, and, indirectly, the intellectual capacity of the father. Each student was asked what kind of work his father did. These jobs were then classified into seven major kinds of occupations, using *The Dictionary of Occupational Titles* (10) as a guide. In Table III one can observe a rather large variation in mean traditional values of students among the various occupational groupings. Boys whose fathers were in professional work had mean traditional values very close to the mean of 31.00 for all 200 boys in the study, whereas the mean traditional score of 32.08 for girls with fathers in professional work is well above the mean of 30.04 for all girls. Boys and girls whose fathers were in clerical professions appear to have developed opposite kinds of values—boys were highly traditional whereas girls were highly emergent. The reason is puzzling, to say the least. A similar situation, though not as extreme, is observed in the service and the skilled occupations. Sons and daughters of farmers have high traditional values, whereas those of unskilled laborers have low traditional values. From these data one might speculate that the occupation of parents does influence the values of children, though no distinct pattern is evident and the influence doesn't always affect boys and girls in a similar manner. However, two exceptions are observed in the occupations of agriculture and unskilled workers, where the values of sons and daughters are similar. The agrarian influence and the stable family often associated with farming may be reflected in the traditional values of their sons and daughters. The insecurity, low income, and low social status associated with the unskilled laborer perhaps contribute to the high emergent values of their children. Obviously, more research is needed before any definite statements can be made about the influence of occupations of parents upon values of their offspring.

In order to see if there was any relation between the values of students and their occupational choices, comparisons were made between mean traditional values of students selecting various vocations. These are recorded in Table IV. It is quite apparent for both boys and girls that the higher on the socio-

TABLE III

Comparison of the Mean Traditional Values of High School Boys and Girls by Occupation of Fathers

Occupation of Father	*N*	*Boys Mean "T" Score*	*N*	*Girls Mean "T" Score*	*N*	*Total Mean "T" Score*
Professional-Managerial	23	30.86	34	32.08	57	31.07
Clerical	23	33.04	16	27.93	39	30.95
Service	10	32.00	15	29.26	25	30.36
Agriculture-Forestry	32	32.03	38	31.39	70	30.92
Skilled	55	31.49	35	29.62	90	30.77
Semi-skilled	12	30.25	11	31.72	23	30.97
Unskilled	26	27.15	34	28.20	60	27.75
Retired-Unemployed	19	30.73	17	29.00	36	29.92

TABLE IV

Comparison of Traditional Values of High School Boys and Girls by Occupations of Their Choice

Occupational Choice of Student	*N*	*Boys Mean "T" Score*	*N*	*Girls Mean "T" Score*	*N*	*Total Mean "T" Score*
Professional-Managerial	76	32.40	72	32.30	148	32.35
Clerical	6	32.50	65	28.30	71	28.66
Service	18	31.16	23	27.82	41	29.29
Agriculture-Forestry	29	30.10	1	27.00	30	30.00
Skilled	27	29.66	–	–	27	29.66
Semi-skilled	2	27.50	2	24.00	4	25.75
Unskilled	8	26.12	–	–	8	26.12
Undecided	34	30.68	37	30.46	71	30.56

economic ladder is the occupational category, the higher the mean traditional values of those selecting the vocation. In all categories the values of the girls are more emergent than for boys. This is in line with an earlier observation that girls in general were less traditional than boys. It can be concluded that a relationship perhaps does exist between occupation selected and the personal values of the students. The lower the vocation is on the occupational ladder, the more emergent are the values of those choosing it.

It is of further interest to compare the numbers of students selecting a certain occupational category (Table IV) with the number of parents employed in each grouping (Table III). Almost three times as many students chose professional kinds of occupations as there were parents employed in this category. The reverse is true in the skilled, semi-skilled, and unskilled occupations. There, only 18.5 per cent of the boys chose either of these categories, although 46.5 per cent of their parents were in one of these groupings. The category that came the closest to matching actual employment with student choice was boys whose fathers were in agriculture and the number selecting agriculture as a life's work. Farming has often been referred to as a self-perpetuating vocation.

Grades earned by students also appear to be related to personal values. The significant F value from the analysis of variance (Table V) demonstrates that the differences in traditional values among the achievement levels were too large to be considered the result of chance alone. Students making the highest grades tended to be those with high traditional values, whereas low-achieving students tended to be those with emergent kinds of values. The non-significant interaction indicates a uniformity in traditional-values scores between the sexes.

It is interesting to compare the grades of boys and girls. There again it is demonstrated that girls get better grades than boys. Three times as many girls as boys earned predominantly A grades. The grades do not follow a normal curve, but they do present a distribution that appears very realistic.

The college preparatory curriculum was very popular in the five high schools in this study. It would be interesting to know if the 43 per cent in this curriculum has increased during the past five years. Also, it is worthy of note in Table VI that more boys than girls are in this major whereas girls dominate in the business-commercial major. The general major is popular with boys, for the reason that those interested in terminal type of education will be in this group. Forty students, 26 boys and 14 girls, did not know which curriculum, if any, they pursued.

Here again, definite relationships are shown between traditional values and the factors on which the data were stratified. For boys, a distinct directional trend is apparent. More traditional boys are in college preparatory programs, and the emergent in the general major. For girls, the most traditional group is also in the college preparatory major whereas the most emergent group is in the business-commercial curriculum. This appears to

TABLE V

Analysis of Variance* of Mean Traditional-Value Scores of Boys and Girls by Academic Achievement

Grades Received	*Boys*		*Girls*		*Total*	
	N	*Mean "T" Score*	*N*	*Mean "T" Score*	*N*	*Mean "T" Score*
Mostly A's	8	36.12	24	33.83	32	34.41
Mostly B's	55	32.09	70	30.03	125	30.94
Mostly C's	109	30.37	89	29.16	198	29.82
Mostly D's	27	29.96	16	29.12	43	29.65

F_s 391 (for grades) = 4.331 Significant at the .05 level.
F_s 391 (interaction) = 0.903 Not significant at the .05 level.

*Adjustments were made for disproportionality among groups.

TABLE VI

Analysis of Variance* of Mean Traditional-Value Scores of Boys and Girls by Curriculum in High School

Curriculum	*Boys*		*Girls*		*Total*	
	N	*Mean "T" Score*	*N*	*Mean "T" Score*	*N*	*Mean "T" Score*
College Preparatory	97	33.31	77	31.70	174	32.60
Business-Commercial	13	30.31	77	27.90	90	28.24
General	64	28.55	32	30.78	96	29.29

F_s354 (for curriculum) = 12.385 Significant at the .05 level.
F_s354 (for interaction) = 2.552 Not significant.

*Adjustments were made for disproportionality among groups.

be in line with the data in Table III, which show that girls whose fathers were in clerical occupations were highly emergent. Also, Table IV shows that girls selecting clerical occupations have emergent-value patterns.

The differences between the mean traditional values of boys and girls in the various curriculum patterns are large enough to be statistically significant. The non-significant interaction infers a uniformity in mean traditional-values scores of boys and girls.

An attempt to arrive at student concern for religion through frequency of church attendance produced some interesting inferences. A review of Table VII shows a distinct directional relationship between mean traditional values and regularity of church attendance. For both boys and girls, the more frequent the church attendance the more traditional were their values. Differences in mean scores among categories are so large that they probably didn't happen by chance alone. There was evidence to support the contention that mean traditional-value scores were uniform between the sexes.

It is quite obvious that girls are more regular in church attendance than boys. When the data were stratified by grade levels, it was found that freshmen were more church-oriented than seniors. Senior boys, as a group, were the most lax in church attendance.

It will be recalled that traditional values were made up of four sub-categories: Puritan morality, individualism, work success, and future-time orientation. Emergent values included sub-categories of relativism, conformity, sociability, and present-time orientation. Of the 128 value items in the *Differential Values Inventory,* 16 pertain to each of the above-named sub-groups. Thus, the range of student scores for any category would be 0-16.

Table VIII shows the mean sub-category scores for various student groupings. Note that each traditional value sub-category is followed by the emergent value that is its opposite. As shown earlier herein, the over-all mean traditional score for all students was 30.52, whereas the over-all mean emergent score was 33.45. In the all-student column (Table VIII) these gross means are divided into the means for each of the eight sub-categories. The L.S.D. (Least Significant Difference) row, near the bottom of the table, shows a number that can be used to determine which means in the column are significantly different at the .05 level of confidence.

It will be noted that the mean score for relativism is significantly larger than its opposite for all groupings. The mean sociability score is significantly larger than is work success in all categories except boys. Likewise, present-time scores are generally larger than future-time orientation. If this sample were true to form, conformity should be significantly larger than individualism. However, the reverse is true. So, here we have a large group that is emergent in three categories and traditional in one. This switch from conformity to individualism is interesting. The explanation might be that students reflect the emphasis in the public schools on recognition of individual differences, or there is a mild rebellion against being dominated by the "gang." Also, schools have renewed emphasis on originality and creativity.

A study of Table VIII will further reveal differences in sub-categories of values of freshmen and seniors, and of boys and girls. When comparing grand means earlier in this report for these groups, no significant differences

TABLE VII

Analysis of Variance* of Mean Traditional-Value Scores of Boys and Girls by Frequency of Church Attendance

Frequency of Church Attendance	*N*	*Boys Mean "T" Score*	*N*	*Girls Mean "T" Score*	*N*	*Total Mean "T" Score*
More often than once a week	14	34.57	24	32.29	38	33.13
Once a week	72	32.54	105	30.30	177	31.21
Once a month	30	29.87	31	29.58	61	29.72
Seldom	84	29.48	39	28.46	123	29.16

$F_{3}91$ (for church attendance) $= 5.105$ Significant at the .05 level.
$F_{3}91$ (for interaction) $= 0.395$ Not significant.

*Adjustments were made for disproportionality among groups.

TABLE VIII

Comparison of Mean Traditional and Emergent Scores Paired by Sub-Groups for Various Categories of Students

Value Categories (Paired)	*Groupings of Students*				*All Students N = 400*
	Freshmen N = 200	*Seniors N = 200*	*Boys N = 200*	*Girls N = 200*	
Puritan Morality	6.69	6.54	6.45	6.78	6.61
Relativism	8.98*	9.01*	8.75*	9.23*	8.99*
Individualism	7.46*#	7.97*	7.67*	7.75*	7.71*
Conformity	7.08#	6.38	6.83	6.63	6.73
Work Success	7.88	7.81	8.36#	7.33	7.84
Sociability	9.02*	8.76*	8.47#	9.30*	8.89*
Future Time	8.12#	8.65	8.53	8.24	8.38
Present Time	8.80*	8.93	8.97	8.76*	8.86*
LSD at .05 level	0.507	0.490	0.502	0.492	0.353
F-Value‡	23.564	35.134	26.350	34.880	55.583

*Significantly larger than paired value at the .05 level.
#Significant difference between particular values of groupings at the .05 level.
‡All F-values are significant at the .05 level.

could be demonstrated. Here it can be observed that freshmen are significantly higher in conformity and significantly lower in individualism and future-time orientation than are seniors. Boys are significantly higher in work success and lower in sociability than are girls. So, while in general there was no change in total scores, there are internal differences that are not reflected in the grand mean because they tend to be compensatory.

The F values listed in the bottom row indicate significant differences in the mean value scores in the column directly above the score. Had these not been significant, use of the least-significant-difference procedure would not have been in order.

Highlights of Study

Analysis of the data presented substantiates the following, essentially the highlights of the study:

1. Mean emergent-value scores were larger than mean traditional-value scores for all groupings of these 400 students. In all groups for senior boys these differences were large enough to be significant at least at the .05 level.

2. Differences were not significant in mean total traditional-value scores between freshmen and seniors or boys and girls. Other groupings of the students showed similar results.

3. Even though gross means showed no significant differences in traditional values between freshmen and seniors, and boys and girls, differences were found in the sub-categories. Freshmen were significantly higher in conformity and significantly lower in individualism and future-time orientation than were seniors. Boys were significantly higher in work success and lower in sociability than girls.

4. When traditional and emergent values were broken into the four sub-groups it was found that the emergent sub-groups of relativism, sociability, and present time were in most instances significantly larger than their opposites, Puritan morality, work success, and present-time orientation. However, individualism, a traditional sub-group, was significantly larger than its emergent opposite, conformity. So, although the gross means showed that traditional values were significantly smaller than emergent, this was true in only three of the four sub-groups making up the total.

5. Whether the student came from a home in which there was both a father and a mother or an atypical home situation appeared not to influence the value patterns of students significantly.

6. The wide divergence among the mean traditional-value scores of students stratified by occupation of father lends evidence to the belief that this factor influences the values of students. However, no clear-cut pattern of influence was apparent.

7. A distinct relationship was apparent between occupational choice and mean traditional values. The nearer the occupational grouping selected was to the professional-managerial group, the more traditional were the values of those selecting it.

8. Students receiving high grades had significantly higher mean traditional values than those receiving lower grades.

9. Curriculum followed in high school is related to student values. Those choosing the college-preparatory curriculum were significantly more traditional than were those in other majors.

10. Students who attended church frequently had significantly higher traditional values than those who seldom attended. Girls were more regular in church attendance than boys, and freshmen attended church more often than seniors. Those with high academic achievement attended church more often than those with low achievement.

Conclusions

This study confirms certain findings of other studies and helps demonstrate that the *Differential Value Inventory* by Prince is an acceptable instrument for measuring student values. It was found that the high-achieving student, the church-attending student, the college-preparatory student and the student preparing for one of the professions, tended to have a value pattern leaning toward the traditional. His opposite, the low-achieving, non-church-attending, general-major student, tended to be more emergent than traditional in his values. In general, mean-values scores did not differ between boys and girls or between freshmen and seniors. It will be interesting to compare the value patterns of these freshmen when they are re-measured in three years.

There is need for extensive study of the value patterns of students starting at the elementary school. More and more research bears out the fact that there is perhaps little general change in the value patterns of students during high school and college. However, not all of this evidence is conclusive. If the schools at all levels are to make their maximum contribution to helping students to interiorize the kinds of values that will enable them to become thinking citizens, then we must be able to identify these values specifically. We must also find when these values should be learned, who can be most influential, and how the learning can be most effectively and efficiently accomplished.

REFERENCES

1. Brown, A. L., Morrison, J., and Couch, G. B., "Influence of Affectional Family Relationships on Character Development," *Journal of Abnormal and Social*

Psychology, 42:422-428, October 1947.

2. Crane, A. R., "Pre-Adolescent Gangs: A Socio-Psychological Interpretation," *Journal of Genetic Psychology*, 86:275-279, June 1955.
3. Getzel, J. W., "Values in School and Society," *The High School in a New Era.* (Chase and Anderson, editors) Chicago: The University of Chicago Press, 1958, pp. 150-155.
4. Havighurst, R. J., Robinson, M. Z., and Dorr, M., "The Development of the Ideal Self in Childhood and Adolescence," *Journal of Educational Research*, 40:241-257, December 1946.
5. Hemming, J., "Moral Development in a Changing Society," *British Journal of Educational Psychology*, 27:84, 1957.
6. Linton, R., "Concepts of Role and Status," *Readings in Social Psychology.* (Swanson, Newcomb Hartley, editors) New York: Henry Holt and Co., 1951, p. 266.
7. Phoenix, P. H., "Values in the Emerging American Civilization," *Teachers College Record*, 61:7:355, April 1960.
8. Prince, R., "A Study of the Relationships Between Individual Values and Administrative Effectiveness in the School Situation," Unpublished doctoral dissertation, University of Chicago, December 1957.
9. Riesman, D., Jacob, P. E., and Sanford, N., *Spotlight on the College Student.* Washington, D.C.: American Council on Education, 1959, p. 4.
10. U. S. Department of Labor, Bureau of Employment Security, *Dictionary of Occupational Titles:* 1949, Vol. II. *Occupational Classification and Industrial Index.*

The Junior College Leadership Program of the University of California, Los Angeles, B. Lamar Johnson, Director, has just released *Faculty Handbooks in California Public Junior Colleges*, the first in a series of reports from the Program. This study of 51 public junior college faculty handbooks was conducted by Dr. Frederick C. Kintzer in the Laboratory in Junior College Administration. The Laboratory, which is being developed as a feature of the UCLA program, operates under a grant from the Kellogg Foundation, and has as its purpose the assembling, organizing, and cataloging of materials on junior college administration. Additional reports will be published from time to time. This 25-page pamphlet can be obtained from the UCLA student store; price $1.00.

XII, No.
ember
ORNIA
SOCIATI